

ابزارهای هوشمند فناوری در محصولات جدید نور

سخنرانی حجت‌الاسلام والمسلمین دکتر حمید شه‌ریاری، رئیس مرکز تحقیقات کامپیوتری علوم اسلامی

در آیین رونمایی از نرم‌افزار مقتل‌الحسین (ع)

پنج‌شنبه ۱۳۹۶/۶/۲۳



بسم الله الرحمن الرحيم. خدمت حضار محترم، مهمانان گرامی و همکاران عزیزمان در مرکز تحقیقات کامپیوتری علوم اسلامی و نیز پژوهشگران مؤسسات تحقیقاتی و آموزشی حاضر در این مراسم، عرض سلام و ادب دارم. همچنین، بسیار متشکرم از مرجع عالی‌قدر شیعه، حضرت آیت‌الله العظمی مکارم شیرازی و دفتر معظم ایشان که با پیام شیوا و گهربار خود، ابراز لطف نمودند و با کلمات گهربار خودشان، ما را در راستای اهداف و اقداماتی که در سال‌های گذشته انجام دادیم و ان‌شاءالله در سال‌های آینده نیز در حال انجام آنها هستیم، تشویق کردند. البته بنا بر این بود که از محضر مبارک معظم‌له نیز استفاده کنیم؛ ولی به دلیل کسالتی که برای ایشان عارض شد، چنین توفیقی از ما سلب شد. امیدواریم به مدد الهی هرچه سریع‌تر بهبودی کامل برای معظم‌له حاصل شود.

جایگاه و توسعه محصولات نور

امروز بنا گذاشتم علاوه بر اینکه خطاب به مهمانان گرامی و عزیزان و پژوهشگران حاضر سخن بگویم، مجموعه همکارانم در مرکز تحقیقات کامپیوتری علوم اسلامی نور را نیز مخاطب خود بدانم.

از وقتی بنده در این مرکز در خدمت شما بودم، چیزی حدود بیست و یک سال می‌گذرد که خودش واقعاً یک عمر است و شهادت می‌دهم در طول این سال‌ها، حداقل نیمی از اوقات مفید خویش را برای مرکز صرف کردم و اگر روزی این توفیق حاصل شود که رجعتی حاصل گردد و ما بتوانیم تصمیمات مجددی راجع به زندگی‌مان بگیریم، بنده همین تصمیمی را می‌گرفتم که تا به امروز گرفته‌ام. امروز می‌خواهم دریچه‌ای را در مقابل شما باز کنم؛ البته این مطالب، برای برخی از دوستان ما که در متن این کارها هستند، تکرار مکررات است؛ بلکه خودشان تولیدکننده این کلمات و مفاهیم هستند و من مسئولیت بازخوانی و بازگویی آنها را بر عهده دارم. در این فرصت کوتاه، می‌خواهم با نگاه راهبردی و کلان به مسئله‌ای که بنده متصدی آن هستم، نگاه کنم؛ یعنی با نگاه کلان به سوی آینده این مرکز بنگرم و آنچه را می‌بینم، تا حدی که برایم مقدور است، برای شما ترسیم نمایم.

از همان روزی که مسئولیت این مرکز را به عهده گرفتم، آرزو و آرمانم تولید نرم‌افزار جامع فقه بود. آن قدر به فقه علاقه داشتم و دارم که آن روز تصمیم گرفته بودم اگر چنین نرم‌افزاری را تولید کردم، دیگر کار اجرایی را کنار بگذارم و مشغول تحقیقات فقهی بشوم. حدود پانزده سال طول کشید تا ما به تولید نرم‌افزار جامع فقه دست یافتیم. روزهای پرتلاطمی را نیز در طول آن پانزده سال

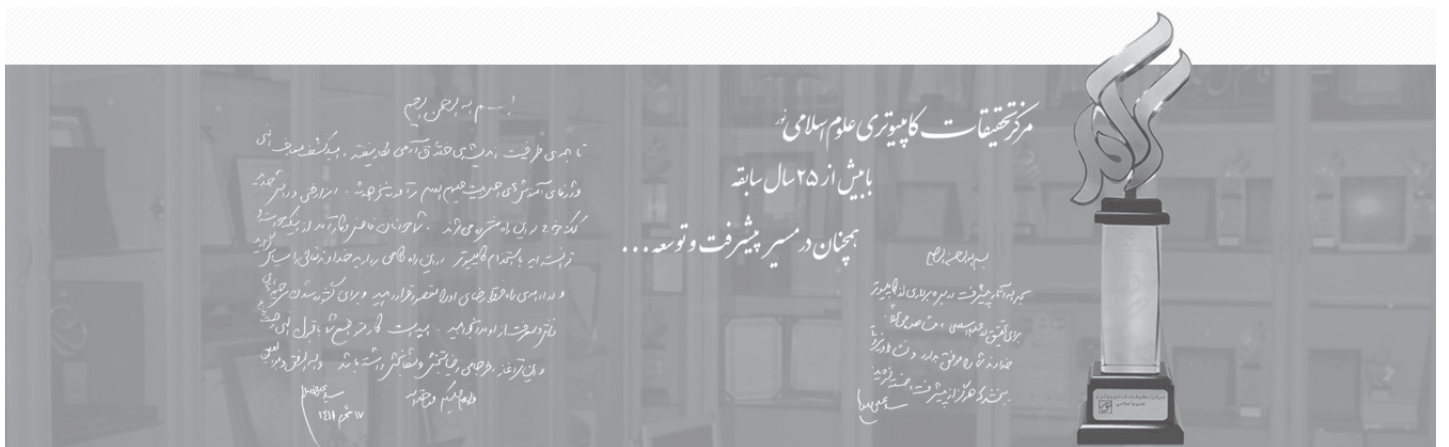
تجربه نمودیم و بسیاری نیز کار ما را تکراری می‌دانستند؛ ولی ما مصمم بودیم که کارمان تکراری نیست؛ بلکه در نوع خود، جدید و بی‌نظیر است.

بحمدالله، امروز می‌توانیم ادعا کنیم مجموعه محصولات نرم‌افزاری نور، فضای آموزشی و پژوهشی در حوزه‌های علمیه را دچار تحول جدی کرده است و خود محققان و طلاب علوم اسلامی، به این معنا اذعان دارند. به برکت این ابزارهای نوین فناوری که در اختیار طلبه‌ها قرار گرفته، همگان به راحتی به منابع فقهی و اسلامی دسترسی دارند و این تحول جدی در نظام آموزشی و پژوهشی حوزه‌های علمیه، به برکت بیست و چند سال خدمت این مرکز و مجموعه‌ای از دست‌اندرکاران آن بوده است. از همین رهگذر است که می‌توانیم ادعا کنیم مرکز تحقیقات کامپیوتری علوم اسلامی، بزرگ‌ترین تولیدکننده دیجیتالی اطلاعات علوم اسلامی در جهان اسلام است؛ یعنی هیچ مرکزی در دنیا به اندازه ما اطلاعات دیجیتالی اسلامی و شیعی در اختیار ندارد و بحمدالله، این مسیر همچنان ادامه نیز دارد.

حرکت به سوی هوشمندسازی نرم‌افزارهای نور

نکته‌ای که در ابتدای عرایضم باید یادآور شوم، این است که فضای فناوری اطلاعات دچار برخی تحولات شده است و اگر فرصت کنم، در فرصتی که در اختیار بنده قرار داده شده، آن را خدمت شما بازگو می‌کنم. بخشی از این تغییرات، به حوزه هوشمندسازی و معنایی مربوط می‌شود و ما حدود هفت سال است که به این مهم رسیده‌ایم و تاکنون فعالیت‌هایی زیرساختی و ارزشمندی را نیز انجام داده‌ایم؛ اما هنوز آن طوری که دلمان می‌خواهد،

امروز می‌توانیم ادعا کنیم مجموعه محصولات نرم‌افزاری نور، فضای آموزشی و پژوهشی در حوزه‌های علمیه را دچار تحول جدی کرده است و خود محققان و طلاب علوم اسلامی، به این معنا اذعان دارند. به برکت این ابزارهای نوین فناوری که در اختیار طلبه‌ها قرار گرفته، همگان به راحتی به منابع فقهی و اسلامی دسترسی دارند و این تحول جدی در نظام آموزشی و پژوهشی حوزه‌های علمیه، به برکت بیست و چند سال خدمت این مرکز و مجموعه‌ای از دست‌اندرکاران آن بوده است



موفق نشده‌ایم همه کارهای خودمان را به محصول قابل عرضه و استفاده تبدیل کنیم و در اختیار پژوهشگران و علاقه‌مندان قرار بدهیم؛ ولی در عین حال، امروز می‌خواهم گزارشی اجمالی از این نوع کارها ارائه دهم که بعضی از آنها عرضه شده و بسیاری نیز هنوز به مرحله عرضه نرسیده است.

به نظرم، حوزه‌های علمیه امروزه نیازمند تحول دیگری هستند که بنده اسم آن را «تحول دوم» می‌گذارم. اگر تحول اول را «آشنایی با فناوری اطلاعات و محتواهای دیجیتالی در محدوده الفاظ» بدانیم، اکنون تحول دوم را باید «دسترسی به محتواهای رقومی از طریق ابزارهای هوشمند و داده‌کاوی معنایی» بنامیم.

در اینجا چند مرحله از فعالیت‌های تشکیلاتی، زیرساختی و نمونه‌هایی از طرح‌ها و اقدامات نرم‌افزاری برای طراحی و تولید ابزارهای هوشمند یا نیمه‌هوشمند را که در گستره دانش‌های حوزوی و زمینه خط و زبان عربی و فارسی انجام داده‌ایم، برای شما عرضه خواهیم کرد.

- ایجاد تشکیلات زیربنایی:

در بخش تشکیلات، چند فعالیت تأسیسی انجام دادیم؛ اول اینکه از سال ۱۳۸۹ گروهی به نام «گروه پژوهشی متن‌کاوی» راه‌اندازی نمودیم که آغازی برای تولید محتواهای دیجیتالی معنامحور و به عبارت دیگر، مقدمه‌ای بر ایجاد سازوکارهای مفهومی در بازیابی محتوا بود. در واقع، تا چند سال پیش در تولیدات نرم‌افزاری، تمرکزمان بیشتر روی بازیابی متون بر اساس الفاظ بوده است؛ اما از سال ۱۳۸۹، به برکت این گروه، تحول چشمگیری در فرایند محصولاتمان شکل گرفت و توجه ما از الفاظ به معنا و محتواها در گستره مفاهیم گرایش پیدا کرد.

دوم اینکه در سال ۱۳۹۱، پژوهشکده فناوری اطلاعات علوم

اسلامی را راه‌اندازی نمودیم و این، اولین پژوهشکده‌ای بود که به اعتبار و همسو با علوم حوزوی و حوزه علمیه تأسیس شد. بسیاری از پژوهشکده‌ها بوده و هستند که اعتبارشان را از دانشگاه‌ها گرفته‌اند؛ یعنی مراجع رسمی دانشگاهی به آنها مجوز فعالیت داده‌اند؛ اما پژوهشکده ما، نخستین پژوهشکده‌ای است که از متن حوزه برخاسته و از شورای عالی حوزه مجوز گرفته است و هم‌اینک نیز در حال توسعه آن هستیم.

سوم آنکه به دنبال تشکیل گروه ادبیات عرب و تحلیل صرفی و نحوی خط و زبان عربی در مرکز قم که از توفیقات خوبی برخوردار بود، گروهی را در معاونت تهران این مرکز، با عنوان «خط و زبان فارسی» به وجود آوردیم که آنها هم از بدو تأسیس این معاونت، مشغول انجام پژوهش‌های کاربردی در این زمینه بودند و بحمدالله از پژوهشگران خوبی نیز در این بخش برخورداریم.

با این سه بخش تشکیلاتی، فعالیت خودمان را در خصوص کارهای مطالعاتی و عملیاتی در بهره‌گیری از هوش مصنوعی شروع کردیم و با استفاده از محتواهای غنی مرکز در موضوعات اسلامی، به مباحث ارزیابی سند روایات به شکل هوشمند و حرکت‌گذاری هوشمند متون عربی ورود کردیم.

- اعراب‌گذاری ماشینی:

حدود پانزده سال، تیم‌هایی از دوستانمان مشغول اعراب‌گذاری روی متون عربی بودند که نتایج خوبی را به دنبال داشت؛ این کار را می‌توان زیربنایی برای اعراب‌گذاری هوشمند ماشینی دانست؛ یعنی امروز می‌توانیم متون عربی بدون اعراب را به شکلی هوشمند و از طریق ماشین، حرکت‌گذاری کنیم. واقعا ابزار حرکت‌گذاری هوشمند متون عربی، کار بزرگی است که در این مرکز به ثمر نهشته است و امروزه، گره خیلی از مشکلات را باز می‌کند. بنده

اجازه ما، چاپ کردند. این اعراب‌گذاری، اول به صورت دستی بود و سپس، الگوهای آن استخراج گردید و به ماشین داده شد و الگوریتم‌های متعددی طراحی شد و پس از بارها تجربه و به‌کارگیری دانش‌های روزآمد، هم‌اینک اعراب‌گذاری به صورت خودکار و توسط رایانه انجام می‌شود. این کار مهم، از نخستین دستاوردهای مرکز محسوب می‌شود. الآن این برنامه‌ها زیاد شده؛ اما دقت کار در برنامه نور، بسیار بالاست و ما اکنون واجد متونی هستیم که دارای اعراب‌های متقن و مطمئن است.

– مشابه‌یابی متون:

از دیگر فعالیت‌های ما در استفاده از تکنیک‌های هوشمند رایانه‌ای، قابلیت مشابه‌یابی لفظی احادیث است که در نرم‌افزار جامع الاحادیث نسخه سه‌ونیم عرضه شده است و در آن، کار مشابهت‌یابی معنایی احادیث انجام می‌شود؛ یعنی می‌توانیم به وسیله نرم‌افزار، الفاظ متفاوت را که هم‌معنا، مترادف یا مرتبط هستند، پیدا کنیم و نمایش دهیم. کار دیگر، شناسایی هوشمند عبارات قرآنی در متون عربی و فارسی است و بیشتر نرم‌افزارهای نور، از این قابلیت مهم برخوردار است که می‌تواند متون قرآنی را استخراج کند و به طور هوشمند نشانی‌دهی نماید.

تقلب‌یابی در خصوص مقالات علمی یا همان کشف سرقت علمی، موضوع مهمی است. متأسفانه، رونوشت‌برداری به نام خود، در برخی مراکز دانشگاهی و حوزوی دیده می‌شود و کشف این موارد، از مباحث لازم است. ما ابزاری تهیه کردیم که می‌تواند میزان متون سرقت‌شده از منابع دیگر را نشان دهد. برای استنادسنجی مقالات و متون علمی با بهره‌گیری از تکنیک‌های جدید برنامه‌نویسی،

می‌توانیم ادعا کنیم مرکز تحقیقات کامپیوتری علوم اسلامی، بزرگ‌ترین تولیدکننده دیجیتال اطلاعات علوم اسلامی در جهان اسلام است؛ یعنی هیچ مرکزی در دنیا به اندازه ما اطلاعات دیجیتال اسلامی و شیعی در اختیار ندارد و بحمدالله، این مسیر همچنان ادامه نیز دارد

شبهه این ابزار را در جاهای دیگر هم دیده‌ام؛ اما اشکالات بسیاری داشته و نمی‌توانسته پاسخگوی نیاز محققان و اهل فن باشد.

رسیدن به الگوی اعراب‌گذاری ماشینی، یک تلاش بیست ساله بوده که از همکاران عزیزم در مرکز تشکر می‌کنم؛ زیرا کار واقعاً طاقت‌فرسایی بوده و الآن یکی از دارایی‌های مرکز، اعرابی است که روی کلمات خورده است. در ابتدای امر، مرکز مجلدات بحار الأنوار را اعراب زد؛ در آن زمان، بحار معرب (اعراب‌گذاری شده) اصلاً وجود نداشت. بعد، برخی از ناشران آن را با اجازه یا بدون

The screenshot shows the website for NOOR Text Mining Group. At the top, there is a search bar with the word 'فارسی' (Farsi) entered. Below the search bar is a navigation menu with items: 'صفحه اصلی' (Home), 'مقالات' (Articles), 'اخبار' (News), 'محصولات' (Products), 'پروژه‌ها' (Projects), and 'نگارخانه' (Gallery). The main content area features a large image with the text 'text mining' and 'noor' overlaid. Below the image, there is a section titled 'پایگاه پژوهشی متن‌کاوی نور راه اندازی شد.' (Noor Text Mining Research Center launched). To the right, there is a 'در باره' (About) section with a paragraph of text. Below this, there are three columns of featured content: 'جدیدترین اخبار' (Latest News), 'پربازدیدترین مقالات' (Most Viewed Articles), and 'محصولات' (Products). Each column contains a list of items with brief descriptions.

شکل مجرور، در یک جمله عربی، چه موقعیت نحوی دارد و چه رابطه‌ای با مستند و مسندالیه دارد. در این صورت است که یک محقق با کشف این رابطه، به مفاهیم متنوعی دست پیدا می‌کند. ما فقط برای دریافت مفاهیم تازه به تجزیه و ترکیب نحوی زبان نیاز نداریم؛ بلکه برای ترجمه ماشینی هم به آن نیازمندیم. امروزه، برخی مراکز هزینه‌های گزافی را برای ترجمه‌های کتب حدیثی و غیرحدیثی انجام می‌دهند. بسیاری از این متون، به دلیل عدم انسجام و یکپارچگی محتواها و عدم اطلاع مترجمان از عبارات و روایات تکراری، دوباره ترجمه می‌شوند؛ درحالی‌که ترجمه ماشینی با دسترسی به بانک داده‌های ترجمه‌شده و بهره‌مندی از هوش مصنوعی، کار ترجمه را آسان و سریع و با کمترین هزینه انجام می‌دهد. اگرچه این کار در اول راه است، اما تلاش مرکز در تکمیل این فعالیت‌ها و فرآیندهای ماشینی، موجب شده به موفقیت‌هایی

نرم‌افزاری تولید کرده‌ایم که به صورت هوشمند این عملیات را انجام می‌دهد. اصولاً یکی از مسائل خیلی مهم که دنیای غرب به آن توجه ویژه دارد و ما کمتر به آن پرداخته‌ایم، استنادسنجی در متون کهن است؛ یعنی هر کدام از دانشمندان متأخر، چه اخذ و اقتباس‌هایی از دانشمندان متقدم خودشان کرده‌اند. این موضوع، برای پژوهشگران تاریخ علم، خیلی مهم است؛ بنابراین، ریشه‌یابی متون فعلی در کتب قدیمی‌تر، بسیار کاربردی و مهم است. در همین راستا، سامانه جست‌وجوگر هوشمند نیز از اهمیت وافری برخوردار است و ما معتقدیم که به طور کلی، ابزارهای هوشمند، کمک وافری به ما خواهد کرد و در واقع، خدمات ارزنده‌ای را به جهان اسلام و آینده علم و دانش ارزانی خواهد داشت.

– نمایه‌زنی ماشینی:

در حال حاضر، ابزارهایی را در دست تولید داریم که امکان استخراج

در سال ۱۳۹۱، پژوهشکده فناوری اطلاعات علوم اسلامی را راه‌اندازی نمودیم و این، اولین پژوهشکده‌ای بود که به اعتبار و همسو با علوم حوزوی و حوزه علمیه تأسیس شد. بسیاری از پژوهشکده‌ها بوده و هستند که اعتبارشان را از دانشگاه‌ها گرفته‌اند؛ یعنی مراجع رسمی دانشگاهی به آنها مجوز فعالیت داده‌اند؛ اما پژوهشکده ما، نخستین پژوهشکده‌ای است که از متن حوزه برخاسته و از شورای عالی حوزه مجوز گرفته است و هم‌اینک نیز در حال توسعه آن هستیم

دست یابد. الآن ما تا حدودی این زحمات و هزینه‌ها را متقبل شده‌ایم و ترجمه‌های مختلف یک متن را در برخی از موضوعات در دسترس قرار داده‌ایم تا محقق بتواند به‌آسانی آنها را بررسی و مقایسه کند. در همین راستا، برچسب‌گذاری صرفی زبان عربی، یعنی تعیین اینکه فلان واژه، «فاعل» است یا «مفعول»، چه نوع مفعولی است، مضاف است یا مضاف‌الیه و یا مواردی از این دست که در بازیابی هوشمند محتوا اهمیت فراوانی دارند، توسط همکاران ما صورت گرفته و حاصل زحمات دوستان، الگویی برای برچسب‌گذاری ماشینی شده است.

– تصحیح املائی خودکار:

از دیگر کارهای ما در این حوزه، تصحیح املائی واژه‌های عربی و فارسی است؛ یعنی شما می‌توانید نگارش خود را تصحیح کنید.

نمایه‌ها را دارد؛ مثلاً نمایه‌زنی ماشینی، از مقوله‌های پیشرفته در حوزه برنامه‌نویسی رایانه‌ای و بهره‌گیری از هوش مصنوعی است.

– تجزیه و ترکیب ماشینی:

از دیگر فعالیت‌های نوآورانه در مرکز نور، تجزیه و ترکیب ماشینی است تا از این طریق، رایانه، تجزیه و ترکیب صرفی و نحوی را به صورت هوشمند انجام دهد؛ یعنی نرم‌افزار می‌تواند عبارات و جملات و کلمات را تشخیص بدهد. بدیهی است که اگر موفق شویم لایه‌های چهارگانه پنهان لفظ را که شامل لایه‌های: صرفی، نحوی، معنایی و گفتمانی است، کشف کنیم، آن وقت می‌توانیم محتوای مورد نظر خویش را سریع‌تر از گذشته استخراج نماییم و به آن دست یابیم و یا به محتواهایی برسیم که ممکن است به ذهن پژوهشگر هم خطور نکند؛ برای مثال، کلمه «عَلِيٌّ» به

این کار، در زبان انگلیسی بیش از بیست و پنج سال است که انجام شده؛ ما نیز امروزه موفق شده‌ایم به تصحیح املائی متن بپردازیم. در خصوص رده‌بندی هوشمند متون و فرهنگ افعال زبان فارسی نیز کارهایی صورت داده‌ایم که فرصت توضیح آنها نیست.

- گنجینه بزرگ علوم اسلامی و انسانی:

به طور کلی، در طی این بیست و چند سال فعالیت مرکز تحقیقات کامپیوتری علوم اسلامی، واجد گنجینه‌ای بزرگ از محتوای دیجیتالی علوم اسلامی و انسانی و به‌ویژه در معارف شیعی شده‌ایم که در مرکز داده نور قرار داده شده و ان‌شاءالله با راه‌اندازی شبکه ملی اطلاعات، قابلیت عرضه در سطح وسیع و گسترده جهانی را دارد. در اینجا اشاره‌ای به این محتواها می‌کنم و آمار مختصری از داده‌های الکترونیکی مرکز نور را ارائه می‌دهم.

تاکنون حدود بیست و دو هزار جلد کتاب را تایپ و تصحیح کرده‌ایم و بیش از چهل و دو هزار جلد کتاب را به صورت دیجیتالی تصویربرداری نموده‌ایم. همچنین، در حوزه مجلات و جراید نیز ده میلیون صفحه مقالات متنی و تصویری را به صورت الکترونیکی در اختیار داریم که از آنها، یک میلیون و صد هزار مجله، به شکل تصویری و بیش از پانصد هزار مجله، در قالب متنی است. همچنین، درصدد هستیم با استفاده از ابزار جدید، متون تصویری خود را به متن تبدیل کنیم. قریب پنج میلیون کلمه غیر تکراری تصحیح شده فارسی و عربی داریم که ریشه دو میلیون کلمه عربی استخراج گردیده و به اصل آن، متصل شده است. این فرایند، در طی پانزده سال صورت گرفته است.

- فرمت‌های تخصصی:

فعالیت دیگر ما، فرمت‌های تخصصی است که متخصصان این مرکز روی کتب: تاریخی، فقهی، اصولی، کلامی، فلسفی و دیگر علوم زده‌اند. همچنین، کدگذاری قریب سیصد و پنجاه هزار روایت، بیش از سیصد و هفتاد هزار سند، بیش از یکصد و بیست و شش هزار کلام غیر معصوم در داخل متن روایات، شصت و هشت هزار فعل یا تقریر معصوم که از متون اسلامی استخراج شده و همچنین، پنجاه و پنج هزار فهرست اعلام کتب و بیست و پنج هزار فهرست اعلام اشخاص، صرفاً بخش کوچکی از فعالیت‌هایی است که در مرکز نور انجام شده و در این فرصت کوتاه، مجال بیان همه آنها نیست.

- استخراج متن از تصویر:

کار دیگری که دوستان ما در حوزه متن‌کاوی انجام داده‌اند، استخراج متن از تصاویر متون است؛ یعنی توانسته‌ایم با الگوریتم

ما ابزاری تهیه کردیم که می‌تواند میزان متون سرقت‌شده از منابع دیگر را نشان دهد. برای استنادسنجی مقالات و متون علمی با بهره‌گیری از تکنیک‌های جدید برنامه‌نویسی، نرم‌افزاری تولید کرده‌ایم که به صورت هوشمند این عملیات را انجام می‌دهد

خاصی، متن تصویری را به متن تایپی تبدیل کنیم. همان‌طور که می‌دانید، تصویر، یک نوع قالب اطلاعاتی است. وقتی شما از یک صفحه متن عکس می‌گیرید، تمام صفحه به یک قلم اطلاعاتی تبدیل می‌شود؛ ولی اگر همان یک صفحه متن را تایپ کردید، به ازای هر کلمه، بلکه به ازای هر حرفی که وارد می‌کنید، یک قلم اطلاعاتی در اختیار دارید و بعداً این اقلام اطلاعاتی، توسط ماشین پردازش می‌شود؛ درحالی‌که اگر شما تصویر داشته باشید، نمی‌توانید چنین پردازش‌هایی را روی آن انجام دهید؛ چون فقط یک قلم اطلاعاتی در اختیار دارید؛ ولی وقتی عکس به متن تبدیل می‌شود، آن وقت است که قابلیت پردازش خواهد یافت. این نیز یکی از کارهایی است که مرکز انجام داده و در حال تکمیل آن است و بحمدالله به نتایج مطلوبی هم رسیده‌ایم.

- موتور مشابه‌یاب متون حدیثی:

همان‌گونه‌که اشاره کردم، موتور مشابه‌یاب حدیث نور در سال ۱۳۹۱ پیاده‌سازی شد و می‌توان گفت یکی از بزرگ‌ترین کارهایی است که مرکز انجام داده است. ما به روش معجم‌های جست‌وجوی لفظی، موتور مشابه‌یاب را طراحی کردیم. این موتور برای اولین بار به همراه قابلیت‌هایی ارزشمند، در نسخه سه‌ونیم نرم‌افزار

در طی بیست و چند سال فعالیت مرکز تحقیقات کامپیوتری علوم اسلامی، واجد گنجینه‌ای بزرگ از محتوای دیجیتالی علوم اسلامی و انسانی و به‌ویژه در معارف شیعی شده‌ایم که در مرکز داده نور قرار داده شده و ان شاء الله با راه‌اندازی شبکه ملی اطلاعات، قابلیت عرضه در سطح وسیع و گسترده جهانی را دارد. در اینجا اشاره‌ای به این محتواها می‌کنم و آمار مختصری از داده‌های الکترونیکی مرکز نور را ارائه می‌دهم

اصل محتوا، تثبیت شده است و به وسیله شاهدهی که توضیح دادم، می‌توان به این توفیق شریف اعتماد پیدا کرد. اینها فواید و برکاتی است که از ابزار مشابهت‌یابی احادیث، نصیب ما می‌شود و کار پژوهشی را جهت‌دهی نموده، به شکلی مطلوبی توسعه می‌دهد.

نظیر این مسئله، در بحث مشابهت‌یابی احادیث فریقین نیز پیاده‌سازی شده است؛ یعنی کتب اربعه شیعه را با صحاح سته اهل سنت مقارنه کردیم و مشابهت‌های مربوطه را استخراج نمودیم و در نرم‌افزاری به نام «جامع احادیث فریقین» قرار دادیم که بسیار مفید و راهگشاست؛ برای مثال، به کمک قابلیت مشابهت‌یابی برنامه، سند «افتتاح الصلاة الوضوء» را از کتاب من لا یحضره الفقیه صدوق نقل کردیم. همین حدیث، با برخی تفاوت‌ها در مسند الدارمی تحت عنوان «مفتاح الصلاة الطهور»

جامع الأحادیث ارائه شد و شما می‌توانید در این برنامه، به دلخواه خود، میزان درصد مشابهت احادیث را تعیین کنید و آنگاه به کشف احادیث مشابه بپردازید که مثال‌ها و نمونه‌های بسیاری را می‌توان برای این امکان کاربردی بیان نمود؛ مثلاً متن توفیق شریف امام زمان (عج) که توسط اسحاق بن یعقوب از دست عثمان بن سعید عمری دریافت شده است و برای مبحثی همچون ولایت فقیه، بسیار کاربرد دارد. یکی از اسناد این حدیث، به محمد بن یعقوب کلینی برمی‌گردد که آن را در کتاب کافی نیاورده است و به همین جهت، برخی این توفیق شریف را مستند نمی‌دانند؛ در حالی که به نظر برخی از بزرگان، این چنین نیست و قطعاً اشکال سندی ندارد.

شما وقتی این متن را با استفاده از قابلیت مشابهت‌یابی احادیث که در نرم‌افزار جامع الأحادیث وجود دارد، مشابهت‌یابی کنید، به حدیثی مثل روایت سعد بن عبدالله از اسحاق بن یعقوب مواجه می‌شوید؛ یعنی برنامه روایتی را نشان می‌داد که یک طرف آن، سعد بن عبدالله اشعری است. وی، از بزرگان قم است که دوران امامت امام حسن عسکری (ع) و دوران غیبت صغری و حضرت صاحب الزمان (عج) را درک کرد. در این حدیث، ایشان از اسحاق بن یعقوب نقل می‌کند. این مطلب، می‌تواند شاهدهی باشد بر اینکه چنین توفیقی، توسط اسحاق بن یعقوب از جانب سعد بن عبدالله اشعری روایت شده است.

بنده وقتی این روایت را مشابهت‌یابی نمودم، مشاهده کردم حدیث مذکور در کتاب الإمامة والتبصرة شیخ صدوق نیز نقل شده است و با کمی تفاوت، در کتاب الخرائج والجرائح قطب راوندی نیز آمده است؛ اما سندی که قطب راوندی ذکر کرده، با آنچه صاحب الإمامة والتبصرة گفته، فرق دارد و در سند ایشان، واسطه اسحاق بن یعقوب نیست؛ بلکه وساطتی است که صحت سند را قطعی می‌سازد و آن، سعد بن عبدالله عن علی بن محمد رازی، معروف به علان رازی است. پس، معلوم می‌شود که



آمده است که باب بحث و تحقیق را برای اهل فن باز می‌کند و نکات جالبی را در معرض فهم و کشف قرار می‌دهد؛ از جمله اینکه به نظر بنده می‌توان به این نتیجه رسید که مثلاً شیخ صدوق، بسیاری از روایات را از منابع اهل سنت گرفته است. به هر حال، اینها نمونه‌های کوچکی است از آثار و نتایج بی‌شماری که این دست ابزارهای هوشمند در معرض استفاده محققان قرار داده است و ما معتقدیم زیرساخت‌های زبانی که سال‌ها روی آنها کار کرده‌ایم، باید سال‌های متمادی دیگر را نیز زحمت بکشیم تا به رشد مطلوب برسد و همان‌گونه که امروز واجد بزرگ‌ترین منابع لفظی در جهان اسلام شده‌ایم، به منابع دربردارنده محتوای معنایی و مفهومی متون اسلامی نیز دست یابیم.

در فایل‌های صوتی برنامه نیز به جست‌وجو بپردازید.

- تحلیل نحوی آیات قرآن:

از دیگر کارهای زیرساخت زبانی که در مرکز انجام شده، تحلیل نحوی آیات قرآن است که شما می‌توانید در محتوای آیات قرآنی به جست‌وجو بپردازید؛ مثلاً وقتی به دنبال آیاتی درباره پدر و مادر بگردید، برنامه تمامی آیاتی را که از نظر لفظی و معنایی به این موضوع اشاره دارد، نشان می‌دهد؛ برای مثال، هم آیه «و وصینا الإنسان بوالدیه إحسانا» را گزارش می‌کند و هم فراز «و صاحبهما فی الدنيا معروفاً» را فهرست می‌نماید که در آن اصلاً لفظ «والدین» نیامده و فقط ضمیر «صاحبهما» به آنها اشاره دارد. این، یک نوع کار محتوایی است که روی آیات انجام شده است.



- موتور جست‌وجوی صوت:

بحمدالله، گام‌های بزرگی در این مرکز برداشته شده که آینده علوم اسلامی را دچار تحولی شگرف خواهد کرد؛ هم‌اینک زیرساخت‌های این مسئله آماده شده؛ اما محصولش هنوز بیرون نیامده است و در گذر زمان، به نتیجه خواهد رسید؛ از جمله آنها، موتور جست‌وجوی صداست؛ یعنی شما از این پس می‌توانید روی صوت‌هایی که در محصولات نور عرضه شده، به جست‌وجو بپردازید؛ مثلاً وقتی کاربر لفظ خاصی را در یک فایل صوتی جست‌وجو کند، نرم‌افزار، او را به آن قسمت از صوت که دربردارنده کلمه مورد نظر است، منتقل می‌کند. این قابلیت، در نرم‌افزار «مقتل الحسین(ع)» که امروز مراسم رونمایی آن را برگزار می‌کنیم، ارائه شده است و شما می‌توانید علاوه بر متون،

- سامانه مشابه‌یابی متون (سمیم نور):

سامانه مشابهت‌یابی متون نور یا سمیم نور، از دیگر تولیدات ماست که مقالات را در گستره بانک غنی مقالاتی که در دسترس ماست، مشابهت‌یابی می‌کند و درصد مشابهت متن مقاله‌ها و پایان‌نامه‌ها و میزان کپی‌برداری آن را از سایر متون نشان می‌دهد. این ابزار، در دانشگاه‌ها، حوزه‌های علمیه و دیگر مراکز تحقیقاتی و علمی، کاربرد بسیاری دارد و بدیهی است که رواج آن در جامعه علمی، عرصه‌های جدیدی را در حوزه پژوهش باز خواهد کرد. این قابلیت در حال حاضر، در پایگاه اینترنتی سمیم نور عرضه شده و برای همگان قابل استفاده است و به راحتی می‌توان با ارائه متن مورد نظر خود، از میزان اعتبار و اصالت آن مطلع شد. برای مثال، شما به کمک این ابزار به آسانی متوجه می‌شوید که چه



بزرگترین بانک مقالات علوم اسلامی و انسانی

– موتور جست و جوگر معنایی:

در اینجا لازم می‌دانم کمی درباره موتور جست‌وجوی معنایی که در برنامه‌ها و سایت‌های مرکز راه‌اندازی شده است، توضیح دهم؛ همچنان که می‌دانید، موتور جست‌وجوی الفاظ در نرم‌افزارهای نور، سال‌هاست که طراحی شده و هر از مدتی، روزآمدسازی شده و اکنون از قوی‌ترین موتورهای کاوشگر لفظی است که در نرم‌افزارهای ایرانی به کار رفته است؛ اما موتور جست‌وجوی معنایی، نیازمند بهره‌گیری از تکنیک‌های پیشرفته و هوشمندسازی برنامه و به‌کارگیری هوش مصنوعی است. در طراحی این موتور، همکاران متخصص و فنی مرکز، تلاش بسیاری کرده و موفقیت‌هایی کسب نموده‌اند که به کاربرد آن در برخی از نرم‌افزارها اشاره کردم.

لازم است در اینجا توضیح دهم که دستیابی به موتور جست‌وجوگر

مقدار از تولیدات بزرگان و نویسندگان اسلامی در قرون گذشته، از آن خودشان است و چه میزان از آن، از کتب یکدیگر نقل شده است؛ البته باید توجه داشت که نقل از کتب دیگران بدون ذکر سند، در زمان گذشته، یک سنت بوده است.

– جست‌وجوی هوشمند در پایگاه نورمگز:

در وب‌گاهی مثل مجلات تخصصی نور نیز قابلیت‌های خوبی در دسترس قرار گرفته و ضمن اینکه شما قادر به جست‌وجوی هوشمند هستید، برنامه نیز در خصوص موضوع تحقیقی‌تان، کلیدواژه‌ها و مقالات مرتبط را نیز پیشنهاد می‌کند و مسیر پژوهش را برایتان آسان می‌سازد؛ مثلاً وقتی موضوع «بیداری اسلامی» را در این پایگاه کاوش می‌کنید، بسیاری از موضوعات مرتبط با این مطلب نیز در دسترس قرار خواهد گرفت.

سامانه مشابهت‌یاب متون نور یا سمیم نور، از دیگر تولیدات ماست که مقالات را در گستره بانک غنی مقالاتی که در دسترس ماست، مشابهت‌یابی می‌کند و درصد مشابهت متن مقاله‌ها و پایان‌نامه‌ها و میزان کپی‌برداری آن را از سایر متون نشان می‌دهد. این ابزار، در دانشگاه‌ها، حوزه‌های علمیه و دیگر مراکز تحقیقاتی و علمی، کاربرد بسیاری دارد و بدیهی است که رواج آن در جامعه علمی، عرصه‌های جدیدی را در حوزه پژوهش باز خواهد کرد. این قابلیت در حال حاضر، در پایگاه اینترنتی سمیم نور عرضه شده و برای همگان قابل استفاده است و به راحتی می‌توان با ارائه متن مورد نظر خود، از میزان اعتبار و اصالت آن مطلع شد



پروژه‌های عظیم و ارزشمندی که در مرکز نور در دست انجام است، به یاری خداوند، به‌زودی نتایج ارزنده‌ای را عاید نظام علمی و پژوهشی کشور خواهد کرد که مایه افتخار جهان اسلام است. امروزه، ما بزرگ‌ترین بانک دیجیتالی اطلاعات اسلامی را در اختیار داریم و تصمیم گرفته‌ایم بزرگ‌ترین اطلاعات دیجیتالی نسخ خطی را نیز جمع‌آوری کنیم تا در این عرصه نیز خدمتگزار باشیم



بتواند دشمن را از غیر دشمن تشخیص دهد و در میدان جنگ، به شکل خودکار با دشمن مبارزه کند. ساخت پهبادها و دیگر ابزارآلات جنگی امروزی نیز از این هوش بهره می‌برد و علاوه بر سطح نظامی، در دیگر حوزه‌ها و علوم و فنون و صنایع نیز کاربرد فراوان دارد.

البته کارهای بزرگ دیگری نیز در مرکز در دست انجام است که بخشی از آنها در معاونت تهران پیگیری می‌شود و متأسفانه، در این جلسه فرصت بازگویی همه آنها نیست و آن را به مجالی دیگر وامی‌نهیم.

پروژه‌های عظیم و ارزشمندی که در مرکز نور در دست انجام است، به یاری خداوند، به‌زودی نتایج ارزنده‌ای را عاید نظام علمی و پژوهشی کشور خواهد کرد که مایه افتخار جهان اسلام است. امروزه، ما بزرگ‌ترین بانک دیجیتالی اطلاعات اسلامی را در اختیار داریم و تصمیم گرفته‌ایم بزرگ‌ترین اطلاعات دیجیتالی نسخ خطی را نیز جمع‌آوری کنیم تا در این عرصه نیز خدمتگزار باشیم.

امیدوارم با همت همکاران و متخصصان ما در این مرکز و نیز مساعدت و همفکری همه شما عزیزان، بتوانیم در آینده نزدیک، آثار بهتر و کارآمدتری را به جامعه علمی، به‌خصوص حوزه‌های علمیه ارائه نماییم و همه این زحمات و فعالیت‌های خالصانه، در جهت رضای حق تعالی باشد و شایستگی آن را بیابیم که از یاران صدیق حضرت ولی عصر(عج) باشیم؛ به برکت صلوات بر محمد و آل محمد. ■

هوشمند چقدر اهمیت دارد. وقتی ما واژه‌ای را در یک محتوا جست‌وجو می‌کنیم، در کاوش لفظی، تنها همان واژه و نهایتاً مشابهات لفظی آن را نمایش می‌دهد؛ ولی در موتور کاوشگر معنایی، نمایه‌ها، عبارات و واژه‌های هم‌معنا و حتی مرتبط با معنا را بازایی می‌کند؛ مثلاً روایتی در مورد روزه‌گرفتن فرزند را به رایانه داده‌ایم: «سُئِلَ عَنِ الصَّبِيِّ مَتَى يَصُومُ؟ قَالَ إِذَا أَطَاقَهُ؛ از امام معصوم(ع) سؤال شد: فرزند ما چه وقت روزه بگیرد؟ امام فرمود: زمانی که طاقت آن را دارد.» موتور جست‌وجوگر مفهومی، روایات مربوط به سنّ و وجوب نماز، طاقت فرزند در روزه‌داری، سنّ احتلام، شرایط تکلیف و مفاهیم مرتبط را نیز فهرست می‌کند؛ درحالی‌که در بسیاری از پاسخ‌ها، نه واژه «صبی» آمده و نه از لفظ «صوم» خبری است. این، به آن جهت است که در طراحی موتور مفهومی، اصلاً به لفظ توجه نمی‌شود؛ بلکه رایانه با تکنیک‌های هوشمند، بُردارهایی ریاضی برای مفاهیم ترسیم می‌کند و هر مفهومی که به بُردار مورد کاوش، نزدیک‌تر و همسان‌تر است، آن را نمایش می‌دهد و به ترتیب ارتباط و درصد نزدیکی به بُردار اصلی، پاسخ‌ها را فهرست می‌نماید. این کار، در حال تکمیل است و به‌زودی نسخه‌های کامل‌شده آن، در محصولات نور عرضه خواهد شد. به نظر ما، چنین کاری بسیار ارزشمند و تحول‌آفرین است و تولید علم را در حوزه‌های علمیه متحول خواهد ساخت.

– تحول در نظام آموزشی و پژوهشی:

در مجموع، همکاران عزیز ما در مرکز تحقیقات کامپیوتری علوم اسلامی، این افتخار را دارند روی زیرساخت‌هایی از این دست کار می‌کنند که تحول‌شگرفی در نظام آموزشی و پژوهشی ما ایجاد خواهد کرد. در واقع، ما داریم با صرف هزینه زیاد، هوش مصنوعی را به حوزه علوم اسلامی وارد می‌کنیم. امروزه در جهان، از هوش مصنوعی در عرصه‌های بی‌شماری استفاده می‌کنند؛ برای مثال، آمریکا در صدد ساخت یک انسان هوشمند مصنوعی است که